



Věc: **Zpráva z pracovní cesty**
Pracovník: **Lenka Kovaříková**
Akce: **Krajská digitalizace (seminář)**
Datum konání: **14. 4. 2017**
Místo konání: **BRNO MZK**

Z hlediska UKN zajímavé:

stav digitalizace v G. Tabulce

Rozšíření programu VISK – Dlouhodobá digitální archivace (Zuzana Kvašková, NK ČR)

- koncepce rozvoje knihoven 2011-2015
- digitalizace v ČR není konsolidovaná, proto pomalý vývoj
- myšlenka vytvoření společného důvěryhodné uložení
- 1. nachystat uložení 2. standardy > v další fázi možné financovat i oblast 1 a 3

Komplexní validátor (Zdeněk Vašek, NK ČR)

zdenek.vasek@nkp.cz

- metodický nástroj, digitalizace v rámci VISK7
- pro kontrolu digitalizovaných dat vytvořených podl NDK studardu (monografie, periodika)
- opensource software
- 2017 pilotní provoz, verze 1.5
- oprava chyb, úpravy pravidel, vauidace strukturálních map, vzaby na PRIMIS, validace dávky
- budoucnost: definování řízených slovníků
- lokální kontrola metadat > validace
- Win, Mac, Linux
- příkazový řádek i grafické rozhraní
- github.com/NLCR/komplexni-validaceotr
- <http://www.ndk.cz/archivace/komplexni-validator>
- <https://github.com/NLCR/komplexni-validator/wiki/instalace>
- kotroluje: strukturální vazby, PREMIS, nekonzistence v datech, poškození dat
- > výstup výpis chyb nad celým balíkem (varování, chyba)
- validace.ndk@nkp.cz
- nástroje: JHOVE, ImageMagick, Kakadu, Jpylyzer > lze jednotlivě vypínat
- validace: obrazu, integrity, bibliografických a technických metadat
- **chyby**
 - mezery - v dublin coru
 - absence povinného elementu
 - chybějící odkaz v metadatech (plný text)
 - dokument ve ztrátové kompresy, místo v bezstrátové

Personální otázky v oblasti digitalizace (Lucie Jirků, SVK HK)

- 7. platová třída – samostatné zajišťování agend souvisejících s ochranou formátů (mikrofilmování, digitalizace, péče o knihovní fondy)

- 11. koordinace digitalizace knihovních fondů na celostátní úrovni, návrhy změn standardů, správa a organizace rozsáhlé digitálních knihoven,...
- řešené
- platová třída – samostatný katalogizátor > tvůrce metadat??
- platová třída – nejčastější
- nové pozice:
 - Správce digitální knihovny
 - Systémový knihovník + Správce digitální knihovny
- vytvořena pracovní skupina

Citace.com – možnost zapojení do Krameria (Zuzana Teplíková, Citace.com)

- citace v katalogu: potřeba kvalitních metadat > možnost propojet s citacePro
- pokud by byl zájem zapojit do Krameria, není problém předělat načítač metadat z MODS

Současný stav ČDK, plánovaný vývoj Krameria (Martin Lhoták, NKA)

<https://github.com> info z branže

Krameries – open source, Solr, shibboleth, Fedora

Česká digitální knihovny = Czech digital library

- zastřešení českých digitálních knihoven
- sklizení metadat
- jedno vyhledávání nad všemi digitalizovanými daty, doporučení pro plné texty
- agregátor pro propojení s Europeanu
- uzavření smluv s agregovanými knihovnami – prozatím jen idexace, právo až pak
- současný problém – sloučení MZK a NDK
- <https://www.czechdigitallibrary.cz/>
- provozuje Knihovna Akademie věd

Rozhraní pro online časopisy

- návrh k realizaci, testování na časopisu Knihy a dějiny
- stránka, pdf, čísla, ročníky, jakým způsobem publikovat v periodiku

ProArc – produkční a archivační systém

- dodržování standardů NK ČR – komplement PSP balíček NDK (METS, MODS, MIX? PREMIS, ALTO, DC)
- přebírání dat z externích systémů (Aleph, Registrdigitalizace.cz)
- podpora URN:NBN, UUID, DOI
- LTP
- Open source

Nástroje pro digitalizaci – novinky (Pavel Kocourek, Search Technologie)

- ProArc 3.2.1.
- svázáno s Fedorou

Registr digitalizace

- www.registrdigitalizace.cz
- algoritmus pro importy, riziko duplicit, kvalita
- **INPROVE** – prostředí pro Průzkum, Ochranu, V7zkum a Evidenci novodobých knihovních dokumentů (popis exempláře, ...)
- **VDK** – konzervační
- **PNF** – evidence průzkumů a měření, sledování fyzického stavu fondů, nomitorování konzervátorských zásahů, průběžná ochrana

DML-CZ (Miroslav Bartošek, ÚVT MU)

- na MUNI 4-5 lidí zaměřující se na vývoj knihovních systémů (150 dalších, které drží IT MUNI)
- vše z matematiky na celém světě
- Matematický ústav Akademie věd – ÚVT MU
- součástí Evropské matematické knihovny (Polsko)
- časopisy, sborníky, monografie, osobnosti
- vyhledávání (search) i procházení (browser)
- **možno nastavit automatické embargo**

Technologie

➤ **metadatový editor**

- kompletní zpracování dokumentu, kontroly, integrace (vlastní systém vyvinutý pro DML-CZ)
- umožní vybudovat hierarchickou strukturu (časopis-ročník-číslo-článek)
- import zdrojových dokumentů (tiff, pdf)
- import metadat (základní, OCR, reference,...)
- sestavení článku (z celého čísla dokáže najít texty článků)
- tvorba metadat,
- soubor autorit
- kontroly a validace
- Dspace vyhledávání – opravy
- > statistiky
- export do Dspace
- Správa uživatel/editorů (kdo na čem, kdy, jak dlouho dělá)
- > kontrolní dokument > do Dspace > stav v jaké dokument je
 - OCR (FineReader, Inty, výpočet podobnosti)
 - DML-CZ workflow (viz obrázek)

Digitalniknihovna.cz (Pavla Rychtářová, Jan Rychtář)

support@digitalniknihovna.cz

- *jednotné uživatelské rozhraní Krameria* (26 přispěvatelů), hlavní potřeby uživatel
- denně cca 700 návštěv (11 minut/uživatel)
- dobře vyhledávané přes google, wiki
- novinky
 - optimalizace prohlížečky – hladký zoom
 - výřez a jeho uložení v jpg (IIIF protokol v emageserveru, zatím 5 knihoven)
 - prokliky v podrobnostech (autoři, klíčová slova, jazyk, typy dokumentů)
 - doplněno hledání (lupa, filtry)
 - přidání zdrojové knihovny dokumentu přímo u popisku knihy
 - uživatelské účty – oblíbené

Zvukové dokumenty z hlediska dlouhodobého uchování: standardy a formáty

(Lenka Damborská, Natálie Ostráková)

- vydané na gramofonové desky
- bude brzy zveřejněn
- zvuková kolekce + strana desky + stopa (celá strana desky/více stop na jedné straně desky) + obal
- metadatový popis
- identifikátory (číslo matrice, objednávací číslo)
- výstupy digitalizace
 - zvukové soubory (stopy)

- původní neupravené zbučkové soubory (SA source audioú, WAV
- archivační kopie (MCA – master copy audio), bez úprav i když je nekvalitní, frekvence 96kHz, bitová hloubka 24
- produkční kopie WAV
- uživatelská kopie (user audio) MP3
- metadata (biblio, technická, administrativní, strukturální)
 - MODS, DC, MIX, AES57, PREMIS, METS (structMap, structLink, amdSec)
- plnění nástrojů Jhove, Fits
- bezztrátová komprese FLAC (testuje se, neví se jak snadno validovat)
- MP3 vyhovuje

Spolupráce mezi knihovnami v oblasti IT (Rudolf Kreibich, NK ČR)

rudolf.kreibich@npk.cz

- nedostatek personálu > přeškolení horníky z ostravska na programátory :D
- vytvoření spolupracující IT komunity (nedostatek úvazků všude)
- SDRUK: Sekce pro informační technologie
 - digitalizace, dlouhodobá uchování, online služby
 - komunikační kanály
 - <https://github.com/ceskaexpedice/kramerus/issues>
 - kramerus@lib.cas.cz
- co se provozuje:
- Kramerus, Fedora, SOLR, Image Server, PostgreSQL, MariaDB, Tomcat, Sentry, Supervisor, uwsgi, Elasticsearch (ELK, Apache, Nginx, Jenkins)
- Docker, LXC, OpenShift, Ansible, Vagrant
- Aleph, Clavius, KOHA Evergreen
- Jakým způsobem začít komunikovat?
- vytvářet workshopy, určit společné priority
- jde přeorganizovat instituce, tak aby se přesunul personál do IT
- získat peníze na personál NAKI?
- Vzdělávání pracovníků (CASLIN, UISK, KISK, Opava), získání nástrojů (Pavel Kocourek)

ARCLib – stav projektu a další vývoj (Zdeněk Hruška, MZK)

<https://arclib.cz>

- NAKI II: 2016-2020
- KNAV, MZK, NK ČR, ÚVT MU + externí odborníci
- cílem open source LTP systém
- Metodika pro bit-level ochranu dat
- Metodika pro logickou ochranu dat
- ve shodě s OAIS (ISO 14721)
- vlastnosti systému: minimalistické (ale funkcí) GUI, API, modularizace, paralelizace
- data management, administrace, validace, zpráva účtů/editorů, archival storage, plán ochrany, prezentace
-

Dlouhodobé uchovávání e-born dokumentů (Ladislav Cubr, NK ČR odbor digitálních fondů)

ladislav.cubr@npk.cz

- el. Publikace vydavatelů, webový obsah, elektronické archiválie, šedá literatura
- e-knihy – el. Obdoba tištěných knih, primárním obsahem je text, 1 soubor
- povinný e-výtisk
- bitová/logická ochrana
- Holandská národní knihovna e-knižní velmoc
- formáty: pevná sazba (PDF, DjVu), plovoucí sazba (MOBI, IBOOK, ePub aj.)

- Kontejnerové formáty
- archivační formáty: PDF/A, ePub - logická ochrana
- Ostatní formáty – bitová ochrana
- **ePub2**
- OCF (Open Container Format) – kontejner (zal. Na forátu ZIP)
- OPS (Open Publication Structure (obsah (XML, JPG? PNG, GIF, SVG))
- OPF (Open Packagign Format) – metadata
- **PDF/A**
- PDF/A-1: 1a, 1b
- PDF/A-2: 2a, 2u, 2b
- PDF/A-3: 3a, 3u, 3b
- - *bez závislosti na externích prvcích (fonty) – převod VŠKP do PDF/A*
- konverze:** vydavatel/archiv, vhodné nástroje, manuální úpravy textu
- validace formátu: veraPDF, projekt PREFORMA, ostrá verze v létě 2017
- identifikátory: ISBN, DOI, URN:NBN
- metadata

Důležité pro UKN

na všech možných místech jede SOLR

formát FOXML

kvalifikační práce ve formátu PDF/A

Dspace je potřeba využívat pouze jako archiv, chtělo by to najít metadatový editor (má DML-CZ)

